

The Baryonic Tully-Fisher Relation cares about the Galaxy Sample

Jenny G. Sorce[★], Quan Guo

Leibniz-Institut für Astrophysik, 14482 Potsdam, Germany

ABSTRACT

The Baryonic Tully-Fisher relation (BTFR) is a clear manifestation of the underlying physics of galaxy formation. As such, it is used to constrain and test galaxy formation and evolution models. Of particular interest, apart from the slope of the relation, is its intrinsic scatter. In this paper, we use the *EAGLE* simulation to study the dependence of the BTFR on the size of the simulated galaxy sample. The huge number of datapoint available in the simulation is indeed not available with current observations. Observational studies that computed the BTFR used various (small) size samples with the only obligation to have galaxies spanning over a large range of masses and rotation rates. Accordingly, to compare observational and theoretical results, we build a large number of various size datasets using the same criterion and derive the BTFR for all of them. Unmistakably, there is an effect of the number of galaxies used to derive the relation. The smaller the number, the larger the standard deviation around the average slope and intrinsic scatter of a given size sample of galaxies. This observation allows us to alleviate the tensions between observational measurements and Λ CDM predictions. Namely, the size of the observational samples adds up to the complexity in comparing observed and simulated relations to discredit or confirm Λ CDM. Similarly, samples, even large, that do not reflect the galaxy distribution give on average biased results. Large size samples reproducing the underlying distribution of galaxies constitute a supplementary necessity to compare efficiently observations and simulations.

Key words: galaxies: spiral - methods: numerical - methods: statistical - galaxies: formation - galaxies: evolution

1 INTRODUCTION

The underlying physics of galaxy formation is visible in the empirical two-or-more galaxy parameter correlations. Among these relations, the Baryonic Tully-Fisher relation (hereafter BTFR, e.g. McGaugh et al. 2000; Bell & de Jong 2001; Gurovich et al. 2004; McGaugh 2005; Pfenniger & Revaz 2005; Begum et al. 2008; Trachternach et al. 2009; Stark et al. 2009; Gurovich et al. 2010; Zaritsky et al. 2014), extension of the Tully-Fisher relation (hereafter TF, e.g. Tully & Fisher 1977; Pierce & Tully 1988; Teerikorpi 1995; Giovanelli et al. 1997b; Courteau & Rix 1999; Tully & Pierce 2000; Giovanelli et al. 1997a; Karachentsev et al. 2002; Bedregal et al. 2006; Noordermeer & Verheijen 2007; Springob et al. 2007; Williams et al. 2010; Tully & Courtois 2012; Mocz et al. 2012; Rawle et al. 2013; Torres-Flores et al. 2013; Sorce et al. 2013, 2014b) to the low luminosity galaxies, is often used to constrain and test galaxy formation and evolution models, with various computational methods, in the Λ CDM scenario as well as in alternative theories (e.g. Steinmetz & Navarro 1999; Mo et al. 1998; McGaugh & de Blok 1998; Eisenstein & Loeb 1995; van den Bosch 2000; Mayer & Moore 2004; Gnedin et al. 2007; Governato et al. 2007; Avila-Reese et al. 2008; McGaugh 2012; Aumer et al. 2013; Dutton 2012; Desmond & Wechsler 2015). While the TF links galaxy luminosities to their rotation rates, the BTFR adds galaxy gas in the equation to correlate galaxy baryonic masses and rotation rates.

The baryonic mass (M_{bar}) of a galaxy is defined as the sum of its gas (M_g) and stellar (M_*) masses.

Of particular interest are the slope and the intrinsic scatter of the BTFR as they impose strong constraints on galaxy formation and evolution models. A variety of slopes and scatters have been measured over the past years (see references above). The various measured slope and scatter values for the BTFR can mostly be reconciled considering that these measurements are affected by the stellar normalization (Zaritsky et al. 2014), by the rotation rate retained measurements (McGaugh 2012) and by the observation uncertainties. The latter, indeed, add up to the measurement (of the galaxy baryonic mass and rotation rate) challenge and to the various methods and calibrations available once the rotation curve is acquired (e.g. Persic & Salucci 1991; Courteau 1997; Courteau & Rix 1999; Verheijen 2001; Noordermeer & Verheijen 2007; Courtois et al. 2009). While some works attempt to reconcile the different BTFRs by using different method-obtained rotation rates (Walter et al. 2008), some focus on the stellar normalization (e.g. Zaritsky et al. 2014) and others disentangle the intrinsic scatter from the total observational scatter of the BTFR (e.g. Gnedin et al. 2007; McGaugh 2012). While it is clear that these three variables (stellar normalization, rotation-rate measurement and observational uncertainties) affect the slope and the scatter of the BTFR and while works dedicated to deriving the BTFR are always careful to use a wide spanning range of spiral galaxy properties (in terms of both rotation rate and mass) not to bias the relation, there is currently no clear attempt known to us to study the impact of the galaxy sample size on the measured slope and intrinsic scatter of the BTFR.

[★] E-mail: jsorce@aip.de

In this paper, we propose, everything else aside, to investigate the impact of the number of galaxies used to derive the BTFR slope and intrinsic scatter, as well as the effect of the galaxy distribution in the samples. For that purpose, we use the *EAGLE* simulation (Schaye et al. 2015a, described briefly hereafter in Section 2) as it represents a fair patch of galaxy distribution in the Universe assuming a Λ CDM cosmological model, ideal for our purposes. In the third section, we study the BTFR using successively all the spiral (defined as such using a selection criterion) galaxies in the simulation and then using reduced size samples and binned samples. In Section 4, we compare our findings with measured BTFR obtained with different datasets in the literature before concluding.

2 THE HYDRODYNAMICAL SIMULATION

The *EAGLE* simulation comprises a set of runs with different box sizes and mass resolutions. Here we use the largest *EAGLE* simulation to have a sufficient number of simulated galaxies into our possession to test effectively the impact of the sample size on the BTFR. This simulation follows $1504^3 \approx 3.4 \times 10^9$ dark matter particles and the same number of gas particles in a comoving 100^3 Mpc^3 cubic volume within the framework of a flat Λ CDM cosmology with parameters taken from the Planck mission (Planck Collaboration et al. 2014): $\Omega_m = 0.307$, $\Omega_\Lambda = 0.693$, $\Omega_b = 0.04825$, $H_0 = 67.77 \text{ km s}^{-1} \text{ Mpc}^{-1}$ (i.e. $h = 0.6777$), $\sigma_8 = 0.8288$ and $n_s = 0.9611$. It has been performed with *GADGET-3* (Springel 2005) modified by a pressure-entropy formulation of SPH (Hopkins 2013), new prescriptions for viscosity and thermal diffusion (appendix B in Schaye et al. 2015b, and Dalla Vecchia in prep.) and a different time stepping scheme (Durier & Dalla Vecchia 2012). A Friends-of-Friends (FoF) algorithm (Davis et al. 1985) allowed to identify the halos, while the *SUBFIND* code (Springel et al. 2001; Dolag et al. 2009) gave the bound structures that these latter encompassed. The galaxies are defined as ensembles of stars within these gravitationally bound sub-haloes.

With the subgrid parameters for feedback calibrated to match the observed $z \sim 0$ galaxy stellar mass function, *EAGLE* simulations are able to reproduce a representative set of observed galaxy properties or relations at low redshift, such as specific star formation rates, Tully-Fisher relation and luminosity functions at $z = 0.1$ from the ultraviolet to the near-infrared bands (Trayford et al. 2015).

Simulated galaxies with u, g, r, i, z magnitude bands (the minimum stellar mass is $10^{8.5} M_\odot$) are selected from *EAGLE* public released halo and galaxy catalogues (McAlpine et al. 2015)¹ at redshift $z = 0$. We define the baryonic mass (M_{bar}) of simulated galaxies as the sum of the gas and stellar mass within the 30 kpc aperture (*Mass.Gas* and *Mass.Star*). Schaye et al. (2015b) showed that this choice of fixed 30 kpc 3D aperture yields the similar stellar mass function as a Petrosian aperture. The luminosities of galaxies with stellar mass $< 10^{11} M_\odot$ are not sensitive to the exact choice of the aperture, since the fraction of light in the extended halo is much lower (Trayford et al. 2015).

As for their rotation rate, we take the maximum value of the circular velocity (V_{max} from the *SubHalo* table), defined as $V_{\text{max}} = G(M(< r)/r)^{1/2}$, where $M(< r)$ is the total mass enclosed in a sphere of physical radius r . We will come back to this choice in Section 4.

We make another cut at the u band magnitude $M_u = -16$ to ensure the completeness of the sample according to the luminosity function following criteria of most of the observational studies. We are interested in having a large sample of spiral galaxies spanning over a large range of properties notably in mass and in rotational velocity, a

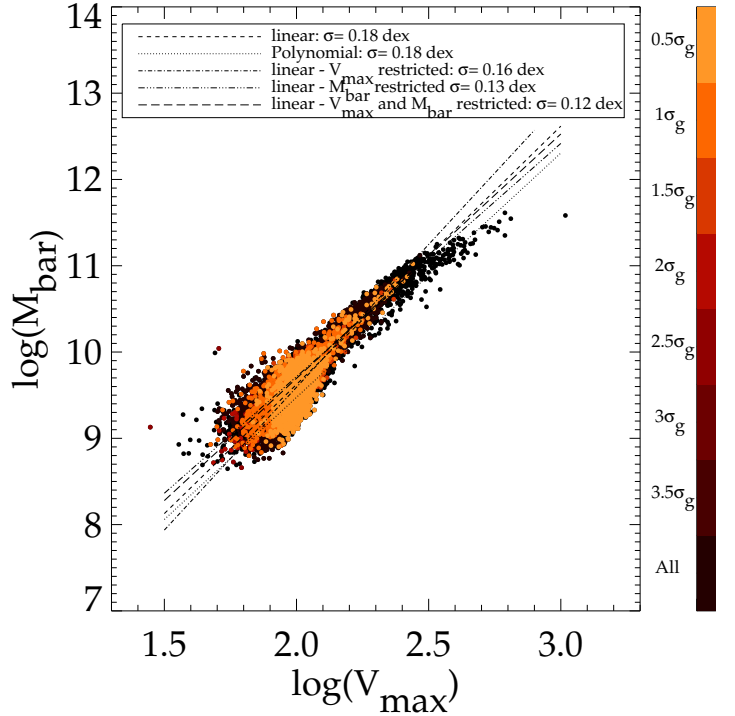


Figure 1. Baryonic Tully-Fisher relation for the sample of spiral (selected using $\log(\text{SFR}/M_*) > -2.0$) galaxies in the *Eagle* simulation (10,225 galaxies) complete down to $M_u = -16$. Dashed and dotted lines are for linear and second order polynomial fits while the dot-dashed, triple-dot-dashed and long-dashed lines fit galaxies with $\log(V_{\text{max}})$ between 1.9 and 2.5, $\log(M_{\text{bar}})$ between 9.5 and 11 and these two restrictions together. The legend gives the (intrinsic) scatter σ around these different fits. The gradient of color shows the distribution of galaxies assuming that latter to be a 3D-Gaussian with a standard deviation σ_g .

cut at a smaller magnitude in absolute value would include irregular dwarfs in the sample while a larger magnitude cut in absolute value would remove the low mass spiral galaxies². Actually, to exclude elliptical galaxies from the sample or conversely to extract spiral galaxies, we use the specific star formation rate (sSFR) to separate star-forming galaxies from passive ones. We apply a cut at $\log(\text{sSFR}) = -2.0$ for galaxies at redshift $z = 0$ following Furlong et al. (2015), who based their choice on a compilation of different observational studies. Galaxies above this cut are considered to be star-forming galaxies and by extension spiral galaxies in the rest of this paper in agreement with observational studies. Indeed, Lee et al. (2007) showed that galaxies with rotational velocities above 120 km s^{-1} are spiral galaxies unless their star formation rate is low (thus they are removed by the selection criterion); galaxies with intermediate rotational rate (between 50 and 120 km s^{-1}) are mostly late-type spirals or alternatively massive irregular that follow the TFR; galaxies with low rotational velocities (below 50 km s^{-1}) are mostly dwarfs galaxies but are in a minority (less than 10) in our sample thanks to the cut in magnitude.

3 THE SIMULATED BTFR

3.1 Of the Total Sample

Schaye et al. (2015a) have already checked the stellar TF relation (correlation between stellar mass M_* and circular velocity V_{max}) for the *EAGLE* simulation. Here we extend their analysis by introducing the gas

¹ <http://www.eaglesim.org/database.html>

² Note that regardless, our main conclusion is unaffected by this magnitude cut.

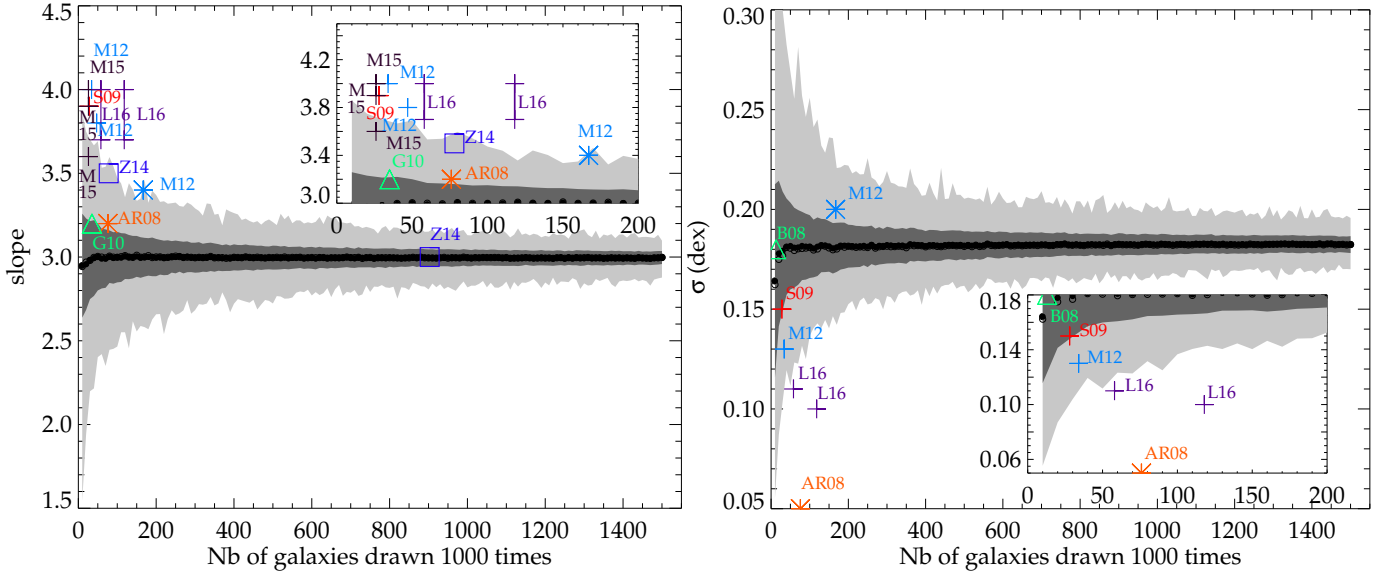


Figure 2. Slope (left) and intrinsic scatter (right) of the simulated BTFR fit as a function of the size of the galaxy sample. Filled and open circles show means and medians (when distinguishable from the former). The dark grey area represents the 1σ scatter of the slopes and intrinsic scatters for each sample size reproduced 1,000 times. The lighter grey area gives the maximum and minimum for each sample size drawn 1,000 times. Colored symbols (one per type of rotational measurements used) stand for results from observational datasets. The small panels are zooms onto the regions where most of the observational values are.

mass (M_g) in the study and we turn to the BTFR. Using the cut described in section 2., we are ‘left’ with a large sample of 10,225 simulated galaxies that abide by the BTFR plotted on Figure 1. One can immediately see that a correlation exists for a large spanning range both in baryonic mass (M_{bar}) and in rotational velocity (V_{max}). We fit two different relations to the datapoint: a linear fit resulting in a 3.0 dex slope and in a 0.18 dex intrinsic (by definition since simulated data have no observational error) scatter. Following Neill et al. (2014), we fit also a second order polynomial and recover the same intrinsic scatter value. However, restricting the linear fit to a smaller range of V_{max} ($80 < V_{max} < 315 \text{ km s}^{-1}$), the scatter is reduced by 0.02 dex and the slope is increased by 0.3 dex. On the contrary, restricting the fit to a smaller range of M_{bar} ($3 \times 10^9 < M_{bar} < 10^{11} \text{ km s}^{-1}$), the slope is decreased by 0.3 dex while the scatter is still decreased but now by 0.05 dex. Restricting both in velocity and baryonic mass decreases the slope by 0.2 dex and the scatter by 0.06 dex. This observation confirms the importance of using datasets with a wide spanning range of properties to grasp the universal BTFR. We checked that these results are not affected by our ‘cut for spirals’ choice as long as the value of that latter stays reasonable.

Before comparing these values to those obtained with observations, it is important to note that currently no observational BTFR study is based on a complete sample and above all none reach the number of galaxies used here. This is understandable as gathering accurate rotation rate and baryonic mass measurements constitutes a great observational challenge. For instance, close to face-on galaxies are almost systematically removed from the samples as it is difficult to determine their rotation rate in such conditions. In the next subsection, we look at the simulated BTFR for various size samples of simulated galaxies.

3.2 Of various-size samples

While the numerous BTFR observational studies currently available never reached the same number of datapoint as the *EAGLE* simulation and while they are never based on complete datasets (a high uncertainty for a galaxy results in its exclusion and there is a limited coverage/overlap of the different required observations), authors are always very careful to use a sample with a large range of velocities and masses

to avoid any bias due to a restricted range of parameters and thus to a non-representative sample of the entire population. However, there is no information about the distribution of these galaxies across the different ranges. We then proceed as follows when building our random selection of varying-size datasets from the simulation: we reject a set when galaxies belonging to that latter do not spread over a wide range of velocities and masses. Namely, we impose that there are galaxies with $V_{max} > 200 \text{ km s}^{-1}$, $V_{max} < 60 \text{ km s}^{-1}$, $M_{bar} < 10^9 M_\odot$, $M_{bar} > 3 \times 10^{10} M_\odot$ (note that varying reasonably these values does not affect the results nor does splitting reasonably M_{bar} over M_* and M_g). We choose these values in agreement with the histograms of V_{max} and M_{bar} to include the tails of these latter. With this selection criterion in mind, simulated galaxy samples containing from 10 to 1,500 galaxies are drawn from the simulations. We repeat the exercise 1,000 times and we fit a linear BTFR relation to all these galaxy datasets. Figure 2 gathers the mean (filled circles), median (open circles, when distinguishable from the former), standard deviation (dark grey area), minimum and maximum (of the 1,000 datasets, light grey area) of all the slopes and intrinsic scatters for each given size sample category.

The result is unmistakable. A too small sample size results in a BTFR slope and an intrinsic scatter that can grandly depend on that very same sample as shown by the large standard deviations around the average BTFR slope and intrinsic scatter. The smaller the sample the larger the probability to measure a different slope from one sample to another. The intrinsic scatter is not spared. Again the standard deviation is very large suggesting that too small samples can give very different results. Table 2 actually gathers the values in percentage chance that a measured slope (scatter) may differ from the universal slope (scatter) at 1σ and 3σ for a given sample size. This table gives a direct access to the uncertainty on a measurement due to the sample size shown on Figure 2. We are able to reproduce these findings with the higher resolution *EAGLE* simulation for the small sample size using the same criteria, however the total number of galaxies is too low to extend the study also to large samples. Regardless, it is reassuring that the larger standard deviations of the slope and intrinsic scatter for the small size samples are not caused by the resolution of the simulation. Before making some comparisons with observational BTFRs from the literature, we look in the next section at the slope and intrinsic scatter

as a function of the circular velocity of the simulated galaxies to understand better how a small sample, although spread over a large range of properties, can give a biased (from the universal relation) result.

3.3 As a function of galaxy parameters

Observations show that the BTFR has a larger scatter for low mass and velocity end (e.g. Begum et al. 2008). In this subsection, we then fit the BTFR for galaxies in different bins of velocities to determine whether this claim is reproduced in the simulation. Slopes and intrinsic scatters as a function of the velocity (bin size of 0.5 dex, plotted only if at least 6 galaxies are in the bin) are given in Figure 3. The left panel of the figure confirms what was already visible in Figure 1, i.e. in agreement with observations, low mass and velocity galaxies are more scattered around the BTFR than their counterparts at higher mass and velocity. The intrinsic scatter (blue filled squares) starts as low as 0.10 dex and increases up to 0.21 dex. Interestingly, the scatter stays quite stable for galaxies with rotation velocities higher than 100 km s^{-1} ($\sigma \approx 0.10$ dex) and increases abruptly for galaxies with rotation rates smaller than 100 km s^{-1} to then reach another plateau. However, processing the galaxies per bin, we are actually restricting the data to small range in velocities and masses, allowing the slopes of the BTFR per bin to vary widely (same panel but black filled circles) from 1.3 to 4 dex.

In the middle panel of Figure 3, we proceed differently by including successively galaxies in the BTFR fit. We start with high rotation velocity galaxies and go down to the low rotation rate ones. The scatter and the slope increase together with the addition of lower and lower rotation rate galaxies. Namely, it confirms that indeed galaxies with a wide spanning range of properties are needed to obtain the slope of the entire sample, however by including more and more low rotation rate galaxies the scatter increases more and more. Again two plateaus are visible with a passage from the lowest one to the highest one at $\log(V_{\text{max}}) \approx 2.0$ dex.

Finally, assuming now that we know the slope of the universal BTFR, we look in the right panel of Figure 3 at the scatter as a function of the rotation velocity with a fixed slope. Again, the scatter is only 0.11-0.12 dex for the high rotation rate galaxies while it rises to 0.2-0.22 dex for the low rotation velocity galaxies.

This implies that although a large spanning range of galaxy properties should in principle enable to get the proper universal slope and intrinsic scatter of the BTFR, it could be that a non-equilibrated sample of galaxies (larger or smaller number of galaxies with high rotation rates with respect to those with low rotation rates) might give a somewhat biased result. We study this hypothesis in the next section.

3.4 As a function of galaxy distributions

To understand the impact of the galaxy sample distribution on the measured BTFR, we vary the criteria applied to keep or reject a dataset to construct Figure 2. Namely, not focusing only on having galaxies spanning over a large range of properties, we observe the following:

- forcing to have approximately the same number of galaxies per 0.5 bin in $\log(V_{\text{max}})$, the average slope varies from 2.9 to 2.7 dex from the small to the large samples and the average scatter varies from 0.15 to 0.19 dex. This implies that an homogeneous distribution across the bins is not adequate to recover the universal BTFR. This is understandable since the galaxies are not homogeneously distributed over the different range of properties.
- forcing to have a higher number of galaxies in the higher bins of $\log(V_{\text{max}})$, the average slope varies from 3.0 to 2.7 dex and the average scatter from 0.145 to 0.175 dex.

- forcing to have a higher number of galaxies in the smaller bins of $\log(V_{\text{max}})$, the average slope varies from 2.9 to 2.7 dex and the average scatter from 0.14 to 0.2 dex.

- forcing to have a distribution of galaxies in the different bins that reproduces the global one, the average slope varies between 3. and 2.95 dex and the scatter from 0.16 to 0.19 dex.

The standard deviations around these average values are still well larger for the smallest size samples than for the largest ones. Still, one can conclude that picking randomly the galaxies ensuring only that galaxy properties span over a large range of values is similar to forcing the distribution of properties in the sample to match that of the entire sample once both are normalized. This is reassuring in the sense that it means selecting randomly the simulated galaxies with the sole constraint of having a wide range of properties, we did not bias the result. However, it is worth noticing that if in a sample there are more high or low velocity galaxies, the sample is biased, even if it is a large sample. It is also important to see that trying to have the same number of galaxies per velocity bins also biases the result. Then the best way to proceed is to pick randomly galaxies provided that this selection is not based on an already biased sample. This is possible for the simulation as all the galaxies are available. This is a priori also possible for observations as galaxies are randomly selected to be observed, although one has to be careful here since observations always favor the brightest galaxies. Then it could be that although galaxies with various properties are constituting an observational sample, this sample is biased not only because it is small but because it has more galaxies in higher or lower (although the first case is more likely than the first one) velocity bins.

This implies that not only the galaxy sample size should be large enough but also the galaxies distribution should reproduce the underlying one. In light of this information, we compare different observational results with those obtained when selecting different size samples spanning over a wide range of properties keeping in mind that for the simulation, it also means reflecting the underlying galaxy distribution.

4 THE OBSERVED BTFR

Table 1 gathers several slopes and intrinsic scatters derived for different observational datasets in the literature. It is important to note that several of these datasets overlap with each others so that we are not expecting too different results between the overlapping sets. Namely, these samples share some identical galaxies and map similar local regions of the sky thus we anticipate that the BTFR measured parameters will lie in comparable areas when overplot onto the simulated results as there are not independently obtained. The observation-derived values are actually plotted in Figure 2 considering the size of the galaxy sample. One can immediately see that unfortunately most of the observational datasets have very often a too small size to be completely trustworthy for comparisons with simulations. They lie in the region where slopes and scatters can vary widely (even with a sample spanning over a wide range of velocity and mass and a galaxy distribution reflecting the underlying one). The sample with the largest number of observed galaxies has on the other hand a slope in perfect agreement with that obtained with simulated galaxies. The intrinsic scatter is unfortunately not directly available but the observational one is 0.29 dex (Zaritsky, private comm.) leaving the possibility of the scatter due to uncertainties to be as high as ~ 0.23 dex (assuming the intrinsic scatter to be 0.18 dex like in the simulation). These values are perfectly reasonable.

On the other hand, observational slopes all (but that obtained with the largest sample) stand on the high end of the values that can be obtained with the simulated galaxy samples. Several issues are probably

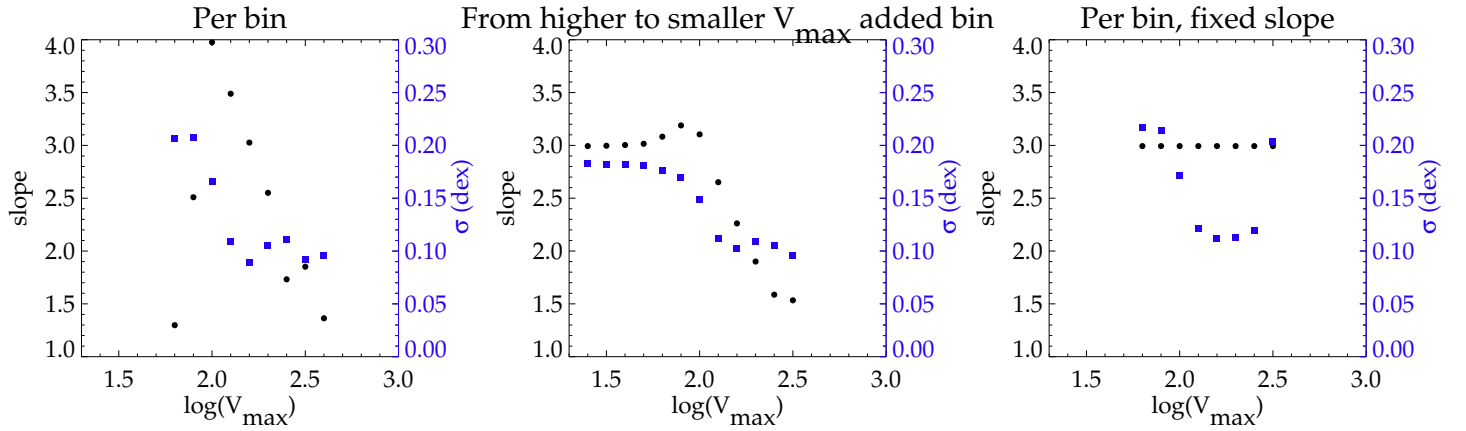


Figure 3. Slope (black filled circle) and intrinsic scatter (blue filled square) of the BTFR as a function of the rotation velocity for simulated galaxies per 0.5 dex bin size (left), per cumulative 0.5 dex bin size, adding from the highest bins in terms of velocity to the smallest ones (middle), per 0.5 dex bin size fixing the slope to that of the total sample (right). A value is shown only if at least 6 galaxies contributed to its derivation.

Nb galaxies	V_{type}	slope	intrinsic scatter	short name	Nb galaxies	V_{type}	slope	intrinsic scatter	short name
13	W_{20}		0.18*	B08	903	W_{50}	3.0		Z14
76	V_m	3.2	0.05	AR08	78	W_{50}	3.5		Z14
28	V_f	3.9	0.15*	S09	26	V_f	3.9		M15
35	W_{20}	3.2		G10	26	V_f	4.0		M15
47	V_f	3.8		M12	26	V_f	3.6		M15
34	V_f	4.0	0.13	M12	58	V_f	3.7-4.0 [†]	0.11	L16
167	W_{20}	3.4	0.2	M12	118	V_f	3.7-4.0 [†]	0.10	L16

Table 1. Properties of the observational datasets when available from the literature. (1) and (6) Number of galaxies in the sample, (2) and (7) type of velocity measurements, either line width W at 20 or 50 % of the HI-profile peak or maximum, m , and flat, f , velocity V (3) and (8) slope of the BTFR, in dex, (4) and (9) intrinsic scatter of the BTFR, in dex, (5) and (10) Short names used in Figure 2. B08 stands for Begum et al. (2008), AR08 for Avila-Reese et al. (2008), S09 for Stark et al. (2009), G10 for Gurovich et al. (2010), M12 for McGaugh (2012), Z14 for Zaritsky et al. (2014), M15 for McGaugh & Schombert (2015) and L16 for Lelli et al. (2016). * these numbers are from McGaugh (2012). [†] range of slope values for stellar normalization coefficients that results in the smaller intrinsic scatter.

in cause: 1) the interdependence of the observed datasets, 2) the cosmic variance of a small local patch of the Universe, 3) the different stellar normalizations, 4) the type of measurements for the rotational velocities, 5) the simulation, based on a model that uses our current (limited) knowledge regarding galaxy formation and evolution.

The matter of the overlapping observational samples has to be addressed directly by observers while that of improving hydrodynamical models is more the responsibility of theorists and developers. The cosmic variance between different small patches of the Universe on the other hand will be reduced in the near future thanks to a combination between hydrodynamical models and constrained simulations of the Local Universe. Indeed, constrained simulations unlike typical cosmological simulations stem from a set of observational constraints thus they resemble the local Large Scale Structure (e.g. Hoffman 2009; Gottlöber et al. 2010; Sorce et al. 2014a, 2016). It is thus expected that combined together with hydrodynamical models, regions similar to the observed one will be reproduced to push the comparisons even further. The slope variation engendered by modifications of the stellar normalization has been discussed to a large extent recently by Lelli et al. (2016). In their paper, the authors show that the slope of the observationally measured BTFR varies between 3.0 to 4.0 dex depending on the stellar normalization coefficient. They settle for coefficients giving slopes in the range 3.7-4.0 that are compatible with a minimum intrinsic scatter of 0.10-0.11 dex but smaller slopes are not excluded.

Finally, there is the delicate question of the rotation rate measurements that differ between observational samples. These distinct

measures result unconditionally in various slopes. In Figure 2, diverse symbols are actually used to distinguish the datasets as a function of the velocity measurements retained to derive the BTFR slope. Flat velocities (crosses) always give larger slopes than linewidth-derived velocities either at 20% (stars) or 50% (squares) of the HI-profile peak or than maximal velocities (triangles). Simulations are not exempt of this issue (e.g. Oman et al. 2016). On top of that, measurements between simulations and observations differ since the dark matter halos are directly available only in the former. Some estimates actually tend to show that the factor between simulated and observed velocity values varies with the mass (Reyes et al. 2012; Katz et al. 2014). Such studies show that simulated velocities should be reduced more on the high end mass than on the low end one to fit observed values. Such a behavior tends to increase slightly the slopes obtained with the simulated samples reducing some more the discrepancy with the observations.

Regardless, overall, observed datasets currently at hands cannot rule out the Λ CDM scenario used to run the simulation. The existing tensions are clearly alleviated when considering, in addition to the already existing concerns mentioned above, also the sample sizes.

Overall, the same ascertainments are valid for the intrinsic scatter. Still, next, we focus some more our attention onto the intrinsic scatter as this is more problematic for the Λ CDM model than the slope that depends more dramatically on the stellar normalization. Lelli et al.

Nb	σ_{slope}		σ_{scatter}		Nb	σ_{slope}		σ_{scatter}		Nb	σ_{slope}		σ_{scatter}		Nb	σ_{slope}		σ_{scatter}	
	in % at		in % at			in % at		in % at			in % at		in % at			in % at		in % at	
	1 σ	3 σ	1 σ	3 σ		1 σ	3 σ	1 σ	3 σ		1 σ	3 σ	1 σ	3 σ		1 σ	3 σ	1 σ	3 σ
10	10	32	30	90	310	2	8	4	14	610	2	6	3	10	910	1	5	2	8
20	9	27	18	56	320	2	8	4	13	620	2	6	3	10	920	1	5	2	8
30	8	25	16	48	330	2	8	4	13	630	2	6	3	10	930	1	5	2	8
40	7	22	13	39	340	2	8	4	13	640	2	6	3	10	940	1	4	2	8
50	6	19	11	35	350	2	8	4	14	650	1	5	3	9	950	1	5	2	8
60	6	18	11	33	360	2	8	4	13	660	1	5	3	10	960	1	4	2	8
70	5	17	10	30	370	2	7	4	13	670	1	5	3	9	970	1	5	2	8
80	5	16	9	29	380	2	7	4	13	680	2	6	3	10	980	1	4	2	8
90	5	15	8	26	390	2	7	4	13	690	1	5	3	10	990	1	4	2	8
100	4	14	8	25	400	2	7	4	13	700	1	5	3	10	1000	1	4	2	8
110	4	14	8	25	410	2	7	4	12	710	1	5	3	9	1010	1	5	2	8
120	4	13	7	23	420	2	7	4	12	720	1	5	3	9	1020	1	4	2	8
130	4	13	7	23	430	2	7	4	12	730	2	6	3	9	1030	1	4	2	8
140	4	12	6	20	440	2	7	4	12	740	1	5	3	9	1040	1	4	2	8
150	4	12	6	20	450	2	7	4	12	750	1	5	3	9	1050	1	4	2	8
160	3	11	6	20	460	2	7	4	12	760	1	5	3	9	1060	1	4	2	7
170	3	11	6	20	470	2	7	4	12	770	1	5	3	9	1070	1	4	2	7
180	3	11	6	18	480	2	7	3	11	780	1	5	3	9	1080	1	4	2	7
190	3	11	6	19	490	2	6	4	12	790	1	5	3	9	1090	1	4	2	8
200	3	11	6	18	500	2	6	3	11	800	1	5	3	9	1100	1	4	2	7
210	3	10	5	16	510	2	6	3	11	810	1	5	3	9	1110	1	4	2	7
220	3	10	5	17	520	2	6	3	11	820	1	5	3	9	1120	1	4	2	7
230	3	10	5	17	530	2	6	3	11	830	1	5	3	9	1130	1	4	2	7
240	3	10	5	16	540	2	6	3	11	840	1	5	3	9	1140	1	4	2	7
250	3	9	5	16	550	2	6	3	11	850	1	5	2	8	1150	1	4	2	7
260	3	9	5	15	560	2	6	3	11	860	1	5	2	8	1160	1	4	2	7
270	3	9	5	15	570	2	6	3	11	870	1	5	2	8	1170	1	4	2	7
280	3	9	5	15	580	2	6	3	10	880	1	5	2	8	1180	1	4	2	7
290	3	9	5	15	590	2	6	3	10	890	1	4	2	8	1190	1	4	2	7
300	2	8	4	14	600	2	6	3	10	900	1	5	2	8	1200	1	4	2	7

Table 2. Values in percentage chance that a measured slope (scatter) may differ from the universal slope (scatter) at 1σ and 3σ for a given sample size. (1), (6), (11), (16) and (21): sample size, number of galaxies ; (2), (7), (12), (17) and (22) 1σ uncertainty on the slope due to the sample size, in % ; (3), (8), (13), (18) and (23) 3σ uncertainty on the slope due to the sample size, in % ; (4), (9), (14), (19) and (24) 1σ uncertainty on the scatter due to the sample size, in % ; (5), (10), (15), (20) and (25) 3σ uncertainty on the scatter due to the sample size, in %

(2016) showed indeed that the intrinsic scatter is also affected by the latter but it is relatively stable and increases only weakly for small normalization coefficients. Interestingly the smaller slope values are also obtained for the smaller coefficients. Typically, to a slope of 3.0 dex corresponds an intrinsic scatter about 0.16-0.17 dex comparable to the values obtained with the simulated samples.

We decide to study in more details the three observational samples that give intrinsic scatters further than 3σ away from that obtained with the same number of simulated galaxies. These are those of L16 (Lelli et al. 2016) and AR08 (Avila-Reese et al. 2008). Even if we have already listed a list of five reasons plus the galaxy sample size has a cause for such an observation, that should really be considered carefully in order to make efficient comparisons in the near future, it is interesting to understand if there exists another culprit, especially since these points are quite below the theoretical predictions.

In Section 3, a hint at the importance of the galaxy distribution across the range of properties has been highlighted. To study in more details the effect, we bin the three above mentioned observational samples and build three sets of 10,000 simulated galaxy samples, each set sharing one of the three property distributions. For the smaller sample of L16, we find an intrinsic scatter of 0.19 ± 0.03 dex (minimum of the 10,000 samples 0.11 dex and maximum 0.31 dex) as shown on Figure 4, top panel. Proceeding similarly for the larger sample of L16, we get an intrinsic scatter of 0.165 ± 0.016 dex (minimum of the 10,000 samples 0.11 dex and maximum 0.25 dex), bottom panel of the same

Figure. For that of AR08, we measure 0.15 ± 0.02 dex (minimum 0.10 dex and maximum 0.24 dex). This clearly shows that a larger sample is not necessarily better than a smaller sample in recovering the universal (for the whole simulated galaxy population) intrinsic scatter: in the three examples, the smallest sample give values on average closer to that obtained for the whole sample although with a larger scatter than for the other two larger samples. AR08-like and large L16-like samples result in low intrinsic scatter on average. Actually, plotting the histograms of properties of AR08 and large L16, we uncover that there are biased towards larger velocity and mass galaxies (i.e. they have more high mass galaxies than low mass ones which is in disagreement with the observed luminosity function), even if they span over a large range of properties. This shows again the importance of not only spanning over a large range of properties but also of reproducing the distribution of the entire population. Regardless, when looking at the AR08 sample, we had some doubts on whether we derive the scatter the same way they did. We believe they proceeded differently. Indeed, taking their datasets by copying datapoint directly from their plot, we obtain in Figure 5 a scatter about 0.20 dex and forcing the slope to be theirs, the scatter increases by 0.01 dex confirming our suspicion of different approaches. On the opposite, we are able to reproduce at 0.01 dex the results from L16 implying that we are in agreement with their way of processing the data. Without knowing in details the way they proceed we simply drop AR08's intrinsic scatter value, while the tensions between the L16 values and the simulated values are allevi-

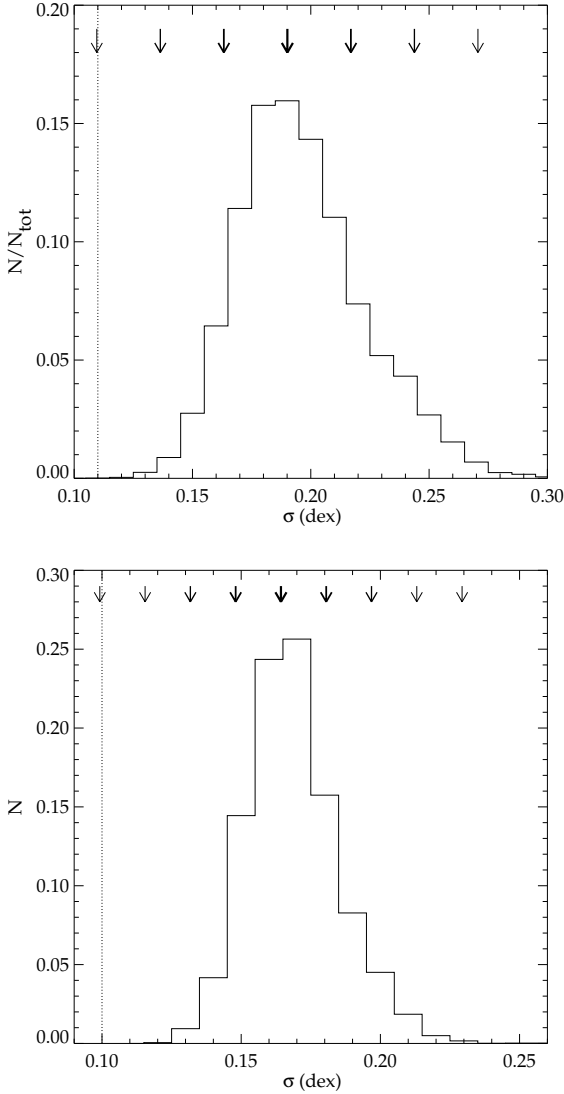


Figure 4. Distributions of the intrinsic scatter found for 10,000 simulated galaxy samples reproducing the property distribution of L16 for the small (top) and large (bottom) samples. Arrows from the thickest to the thinnest show the mean and the 1, 2 and 3, etc σ . The dotted lines stand for the observationally measured values.

ated some more considering also the distribution of galaxies across the range of properties.

5 CONCLUSION

The correlation between galaxy baryonic mass and rotational velocity, or Baryonic Tully-Fisher relation (BTFR), is a clear proof of the underlying physics of galaxy formation. As such, it is often used to constrain and test galaxy formation and evolution models. Of particular interest apart from the slope of the relation is its intrinsic scatter. In this paper, we have studied the BTFR in the *EAGLE* simulation. We measure a slope of 3.0 dex and an intrinsic scatter of 0.18 dex for the entire population of spiral (selected according to a criterion based on the star formation rate and stellar mass, we find that the result is stable for different cuts provided that the chosen value of that latter is reasonable) galaxies. Following observational works that, although they never currently

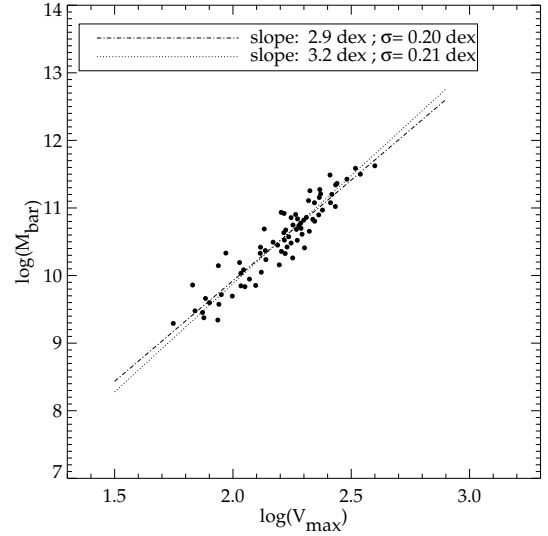


Figure 5. Baryonic Tully-Fisher relation for the dataset from AR08. The dotted dashed line gives the linear fit with all the parameters free, while the dotted line shows the fit with a fixed slope.

have complete and large galaxy samples (mostly because of observational uncertainties and lack of observations), affirm that it is essential to have samples with a wide spanning range in velocity and mass, we build several size datasets using this sole criterion. There is unmistakably an effect of the sample size on the measured BTFR. The smaller the number of galaxies used to derive the BTFR, the larger the standard deviation around the average slope and intrinsic scatter obtained for a given size sample. A sufficiently large sample (about at least 200 galaxies but the larger the better) of galaxies is requested to obtain a trustworthy, in the sense close to the universal, BTFR.

Looking for a reason to that observation, we bin the entire population of simulated galaxies in circular velocity and we notice that like in observations low mass and velocity galaxies present a higher scatter around the BTFR. Actually, repeating the same work as before, i.e. building different size datasets, but forcing not only galaxy to span over a wide range of properties but also to be evenly distributed across bins, or to be over-represented in the higher velocity bins or on the contrary into the smaller velocity bins, we discover that even large samples are biased and do not allow to recover the universal (in the sense from the entire simulated population) BTFR relation. However, forcing the sample to reproduce the overall underlying distribution of galaxy properties, we are able to find back the same result as when picking galaxies randomly requesting only from them to cover a large range of properties, i.e. a non biased result for large samples of galaxies.

Comparing these findings with observation-measured slope and intrinsic scatter values, we immediately notice that a concern arises regarding the dataset sizes. This problem adds up to the already existing list including 1) the interdependence of the observed datasets, 2) the cosmic variance of a small local patch of the Universe, 3) the different stellar normalizations, 4) the type of measurements for the rotational velocities, 5) the simulation, based on a model that uses our current (limited) knowledge regarding galaxy formation and evolution.

Regardless of the fact that observational slopes all (but that obtained with the largest sample) stand on the high end of the values that can be obtained with the simulated galaxy sample probably because of the long list of issues mentioned above, it is clear that the sample size should also be considered to compare efficiently observed and predicted BTFR values. Currently, datasets at hands are small

enough that existing tensions between observations and the Λ CDM cosmological model are alleviated. Focusing on the intrinsic scatter, as it is a powerful constraint for the Λ CDM model, we uncover that all but three observation-derived values are well below the simulation-obtained values (i.e. not within 3σ). Following the hint that results are biased even for large sample if the galaxy distribution does not reflect the underlying one, we decide to look in more details at these three samples of galaxies. We produce datasets of the same size, a similar spanning range of characteristics and distribution across the property ranges. We show that indeed distribution property-like simulated datasets of the same size results on average in smaller intrinsic scatters than those expected for non distribution property-like, but still spanning over a large range of characteristics, samples of the same size.

In summary, the slope and intrinsic scatter of the BTFR depends on a multitude of parameters that all need to be controlled to test efficiently the Λ CDM cosmological model. To the list that already includes for instance the rotation rate measurement, the stellar normalization and the independence of the datasets, this paper permits to add the size of the sample and the distribution of galaxies that should not only span over a large range of properties but should also reflect the underlying distribution of galaxies in the Universe. We are then able to alleviate the existing tensions between observation-measured values and the Λ CDM cosmology. We advocate that large size samples not only spanning over a wide range of properties but also reproducing the overall true distribution of galaxy properties are necessary to envisage pushing further the comparisons with simulations. This highlights the importance of dedicating more observing time in producing larger datasets to study the BTFR.

ACKNOWLEDGEMENTS

JS acknowledges support from the Alexander von Humboldt Foundation. We thank the referee, Stacy McGaugh, for his comments that helped improved substantially the manuscript. We acknowledge the Virgo Consortium for making their simulation data available. The *EAGLE* simulations were performed using the DiRAC-2 facility at Durham, managed by the ICC, and the PRACE facility Curie based in France at TGCC, CEA, Bruyères-le-Châtel. We thank Dennis Zaritsky for providing us with BTFR values.

REFERENCES

- Aumer M., White S. D. M., Naab T., Scannapieco C., 2013, *MNRAS*, 434, 3142
- Avila-Reese V., Zavala J., Firmani C., Hernández-Toledo H. M., 2008, *AJ*, 136, 1340
- Bedregal A. G., Aragón-Salamanca A., Merrifield M. R., 2006, *MNRAS*, 373, 1125
- Begum A., Chengalur J. N., Karachentsev I. D., Sharina M. E., 2008, *MNRAS*, 386, 138
- Bell E. F., de Jong R. S., 2001, *ApJ*, 550, 212
- Courteau S., 1997, *AJ*, 114, 2402
- Courteau S., Rix H.-W., 1999, *ApJ*, 513, 561
- Courtois H. M., Tully R. B., Fisher J. R., Bonhomme N., Zavodny M., Barnes A., 2009, *AJ*, 138, 1938
- Davis M., Efstathiou G., Frenk C. S., White S. D. M., 1985, *ApJ*, 292, 371
- Desmond H., Wechsler R. H., 2015, *MNRAS*, 454, 322
- Dolag K., Borgani S., Murante G., Springel V., 2009, *MNRAS*, 399, 497
- Durier F., Dalla Vecchia C., 2012, *MNRAS*, 419, 465
- Dutton A. A., 2012, *MNRAS*, 424, 3123
- Eisenstein D. J., Loeb A., 1995, *ApJ*, 439, 520
- Furlong M. et al., 2015, *MNRAS*, 450, 4486
- Giovanelli R., Haynes M. P., Herter T., Vogt N. P., da Costa L. N., Freudling W., Salzer J. J., Wegner G., 1997a, *AJ*, 113, 53
- Giovanelli R., Haynes M. P., Herter T., Vogt N. P., Wegner G., Salzer J. J., da Costa L. N., Freudling W., 1997b, *AJ*, 113, 22
- Gnedin O. Y., Weinberg D. H., Pizagno J., Prada F., Rix H.-W., 2007, *ApJ*, 671, 1115
- Gottlöber S., Hoffman Y., Yepes G., 2010, *ArXiv e-prints*: 1005.2687
- Governato F., Willman B., Mayer L., Brooks A., Stinson G., Valenzuela O., Wadsley J., Quinn T., 2007, *MNRAS*, 374, 1479
- Gurovich S., Freeman K., Jerjen H., Staveley-Smith L., Puerari I., 2010, *AJ*, 140, 663
- Gurovich S., McGaugh S. S., Freeman K. C., Jerjen H., Staveley-Smith L., De Blok W. J. G., 2004, *PASA*, 21, 412
- Hoffman Y., 2009, in *Lecture Notes in Physics*, Berlin Springer Verlag, Vol. 665, Data Analysis in Cosmology, Martínez V. J., Saar E., Martínez-González E., Pons-Bordería M.-J., eds., pp. 565–583
- Hopkins P. F., 2013, *MNRAS*, 428, 2840
- Karachentsev I. D., Mitronova S. N., Karachentseva V. E., Kudrya Y. N., Jarrett T. H., 2002, *A&A*, 396, 431
- Katz H., McGaugh S. S., Sellwood J. A., de Blok W. J. G., 2014, *MNRAS*, 439, 1897
- Lee J. C., Kennicutt R. C., Funes, José G. S. J., Sakai S., Akiyama S., 2007, *ApJ*, 671, L113
- Lelli F., McGaugh S. S., Schombert J. M., 2016, *ApJ*, 816, L14
- Mayer L., Moore B., 2004, *MNRAS*, 354, 477
- McAlpine S. et al., 2015, *ArXiv e-prints*: 1510.01320
- McGaugh S. S., 2005, *ApJ*, 632, 859
- McGaugh S. S., 2012, *AJ*, 143, 40
- McGaugh S. S., de Blok W. J. G., 1998, *ApJ*, 499, 41
- McGaugh S. S., Schombert J. M., 2015, *ApJ*, 802, 18
- McGaugh S. S., Schombert J. M., Bothun G. D., de Blok W. J. G., 2000, *ApJ*, 533, L99
- Mo H. J., Mao S., White S. D. M., 1998, *MNRAS*, 295, 319
- Mocz P., Green A., Malacari M., Glazebrook K., 2012, *MNRAS*, 425, 296
- Neill J. D., Seibert M., Tully R. B., Courtois H., Sorce J. G., Jarrett T. H., Scowcroft V., Masci F. J., 2014, *ApJ*, 792, 129
- Noordermeer E., Verheijen M. A. W., 2007, *MNRAS*, 381, 1463
- Oman K. A., Navarro J. F., Sales L. V., Fattahi A., Frenk C. S., Sawala T., Schaller M., White S. D. M., 2016, *ArXiv e-prints*: 1601.01026
- Persic M., Salucci P., 1991, *ApJ*, 368, 60
- Pfenniger D., Revaz Y., 2005, *A&A*, 431, 511
- Pierce M. J., Tully R. B., 1988, *ApJ*, 330, 579
- Planck Collaboration et al., 2014, *A&A*, 571, A16
- Rawle T. D., Lucey J. R., Smith R. J., Head J. T. C. G., 2013, *MNRAS*, 433, 2667
- Reyes R., Mandelbaum R., Gunn J. E., Nakajima R., Seljak U., Hirata C. M., 2012, *MNRAS*, 425, 2610
- Schaye J. et al., 2015a, *MNRAS*, 446, 521
- Schaye J. et al., 2015b, *MNRAS*, 446, 521
- Sorce J. G., Courtois H. M., Gottlöber S., Hoffman Y., Tully R. B., 2014a, *MNRAS*, 437, 3586
- Sorce J. G. et al., 2013, *ApJ*, 765, 94
- Sorce J. G. et al., 2016, *MNRAS*, 455, 2078
- Sorce J. G., Tully R. B., Courtois H. M., Jarrett T. H., Neill J. D., Shaya E. J., 2014b, *MNRAS*, 444, 527
- Springel V., 2005, *MNRAS*, 364, 1105
- Springel V., White S. D. M., Tormen G., Kauffmann G., 2001, *MNRAS*, 328, 726
- Springob C. M., Masters K. L., Haynes M. P., Giovanelli R., Marinoni C., 2007, *ApJS*, 172, 599

- Stark D. V., McGaugh S. S., Swaters R. A., 2009, *AJ*, 138, 392
- Steinmetz M., Navarro J. F., 1999, *ApJ*, 513, 555
- Teerikorpi P., 1995, *Astrophysical Letters and Communications*, 31, 263
- Torres-Flores S., Mendes de Oliveira C., Plana H., Amram P., Epinat B., 2013, *MNRAS*, 432, 3085
- Trachternach C., de Blok W. J. G., McGaugh S. S., van der Hulst J. M., Dettmar R.-J., 2009, *A&A*, 505, 577
- Trayford J. W. et al., 2015, *MNRAS*, 452, 2879
- Tully R. B., Courtois H. M., 2012, *ApJ*, 749, 78
- Tully R. B., Fisher J. R., 1977, *A&A*, 54, 661
- Tully R. B., Pierce M. J., 2000, *ApJ*, 533, 744
- van den Bosch F. C., 2000, *ApJ*, 530, 177
- Verheijen M. A. W., 2001, *ApJ*, 563, 694
- Walter F., Brinks E., de Blok W. J. G., Bigiel F., Kennicutt, Jr. R. C., Thornley M. D., Leroy A., 2008, *AJ*, 136, 2563
- Williams M. J., Bureau M., Cappellari M., 2010, *MNRAS*, 409, 1330
- Zaritsky D. et al., 2014, *AJ*, 147, 134